The Lovasz Bound and Some Generalizations

R. J. McEliece, E. R. Rodemich, and H. C. Rumsey, Jr.
Communications Systems Research Section

In 1956, Shannon defined the zero-error capacity of a discrete memoryless channel as the largest rate at which information can be transmitted over the channel with zero error probability. He exhibited one particularly interesting channel with five inputs and outputs whose zero error capacity he could not compute. The problem of computing this capacity remained unsolved until very recently, when Lovasz computed it in an astonishing simple manner. We show that Lovasz' ideas, combined with some of our own, lead to an extremely powerful and general technique, which we phrase in terms of graph theory, for studying combinatorial packing problems. In particular, Delsarte's linear programming bound for cliques in association schemes appear as a special case of the Lovasz bound.

I. Introduction

Let $V = \{v_1, \cdots, v_N\}$ be a finite set with N elements, and let E be a collection of two-element subsets of V. Then the set G consisting of the singletons $\{v_i\}$ from V and the elements of E is called a *graph*[1] *on* V. The elements of V are called the *vertices*, and those of E, the *edges*, of G. Figure 1 depicts a particularly interesting graph for our purposes:

$$G = \{\{0\},\{1\},\{2\},\{3\},\{4\},\{0,1\},\{1,2\},\{2,3\},\{3,4\},\{4,0\}\}$$

Here vertices are represented by points in the plane, and edges by lines joining appropriate pairs of vertices. (For future reference we label this graph C_5.)

A subset $Y = \{y_1, \cdots, y_M\}$ of V is called an *independent set* if none of the pairs $\{y_i, y_j\}$, $i \neq j$, are edges of G. The cardinality of the largest possible independent set in G is denoted by $\alpha(G)$:

$$\alpha(G) = \max\ \{|Y| : Y \text{ independent set in } G\} \qquad (1)$$

For example, $\alpha(C_5) = 2$, and the set $Y = \{0, 2\}$, circled in Fig. 1, is a maximal independent set.

For any integer $n \geqslant 2$, we now define the n-th *direct power* of G, denoted by G^n, as follows: the vertex set of G^n is the Cartesian power V^n, i.e., the set of all N^n n-tuples $\mathbf{v} = (v_1, \cdots, v_n)$ from V. The edge set of G^n consists of all pairs $\{\mathbf{v}, \mathbf{v}'\}$ from V^n such that $\{v_i, v_i'\} \in G$ for all i. Note that if $\mathbf{A}$ is the incidence matrix of G, i.e., the following $N \times N$ matrix whose rows and columns are indexed by elements of V:

$$\mathbf{A}(v, v') = \begin{cases} 1 \text{ if } \{v, v'\} \in G \\ 0 \text{ if not} \end{cases}$$

[1]More accurately, an undirected graph.

then the incidence matrix of G^n is the n-th direct (Kronecker) power of $\mathbf{A}$.

The *capacity* of G, denoted by $\theta(G)$, is now defined as follows:

$$\theta(G) = \sup_n \alpha(G^n)^{1/n} \tag{2}$$

This notion[2] was introduced by Shannon (Ref. 8) in connection with the problem of finding the zero-error capacity of a discrete memoryless channel, and he developed techniques that enabled him to compute the capacities of many, but not all, graphs.

For example, Shannon showed that if there exists a mapping ϕ of V into an independent set of G such that $\{v, v'\} \notin G$ implies $\{\phi(v), \phi(v')\} \notin G$, then $\theta(G) = \alpha(G)$.

Also, he gave a *linear programming* upper bound on $\theta(G)$, as follows. Let P be a probability distribution on V, i.e., $P(v) \geqslant 0$, $\Sigma \ \{P(v) : v \in V\} = 1$. P is extended to subsets $X \subset V$ additively: $P(X) = \Sigma \ \{P(x): x \in X\}$. The subset X is a *clique* in G if $\{x, x'\} \in G$ for all $x, x' \in X$. Then Shannon proved that

$$\theta(G) \leqslant \lambda^{-1}$$

where

$$\lambda = \min_P \max \ \{P(X): X \text{ a clique}\} \tag{3}$$

the minimization in Eq. (3) being taken over all possible probability assignments.

These two results enabled Shannon to calculate the capacity of all graphs with five or fewer vertices, with the single exception of the graph C_5 of Figure 1. For C_5 his results yielded only

$$\sqrt{5} \leqslant \theta(C_5) \leqslant 5/2 \quad \text{(footnote 3)} \tag{4}$$

Then, twenty-one years later, Lovasz (Ref. 5) established that $\theta(C_5) \leqslant \sqrt{5}$; this, combined with the lower bound in Eq. (4), shows that $\theta(C_5) = \sqrt{5}$. Let us briefly sketch Lovasz' technique for finding upper bounds on $\theta(G)$.

Lovasz defines an *orthonormal representation* of G as a realization of G in which the vertex set V is a set of vectors in a Euclidean vector space, with the property that $\{\mathbf{v}, \mathbf{v}'\} \in G$ iff $\mathbf{v} \cdot \mathbf{v}' \neq 0$.

If G has such an orthonormal representation, and if $\mathbf{b}$ is any unit vector, Lovasz' bound is as follows:

$$\theta(G) \leqslant (\min \ \{(\mathbf{v} \cdot \mathbf{b})^2 : \mathbf{v} \in V\})^{-1} \tag{5}$$

Lovasz applied Eq. (5) to the graph C_5 by considering an "umbrella" with five ribs $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_5\}$ of unit length. If the umbrella is opened to the point where the angle between alternate ribs is 90 deg, then $\{\mathbf{v}_1, \cdots, \mathbf{v}_5\}$ is an orthonormal representation of G in Euclidean 3-space. If the handle $\mathbf{b}$ is also a unit vector, then one easily shows that $\mathbf{b} \cdot \mathbf{v}_i = 5^{-1/4}$ for all i, and hence by Eq. (5), $\theta(C_5) \leqslant \sqrt{5}$.

Lovasz also derived many other consequences of Eq. (5) that we cannot summarize here. However, let us at least remark that several of the examples in Section IV also appear in Lovasz' paper, and were clearly derived by him earlier. (We will give references to Theorems in Lovasz' paper at the appropriate places in Section IV.)

The present paper arose from an attempt to put Lovasz' results into a general setting. We believe we have succeeded in doing this, but in our development orthonormal representations have entirely disappeared. Nevertheless, the bounds we shall derive (at least the bounds on $\theta(G)$) could all be derived from orthogonality graphs, and so we call these bounds *Lovasz bounds*. (Appendix A contains a proof of the equivalence of our methods and Lovasz'.)

In Section II, we give out derivation of the Lovasz bounds. It will be seen that computing these bounds for a fixed graph G amounts to solving a certain nonlinear programming problem.

In Section III, we demonstrate that if the graph G is highly symmetric (in a sense made precise there), this nonlinear programming problem becomes a *linear* programming problem.

In Section IV, we apply our general results to several examples. First we give a very simple bound on $\theta(G)$, which applies to any regular graph. Next, we compute the Lovasz bound for two infinite families of graphs, the cyclic graphs C_N

[2] Actually, Shannon's definition of capacity is the logarithm of our definition.

[3] The lower bound in Eq. (4) results from the fact that $\alpha(C_5^2) = 5$, a fact also established by Shannon.

and the quadratic residue graphs Q_p, which are two different generalizations of the graph C_5. We do not succeed in computing the capacity of any of the graphs C_N with odd $N \geqslant 7$, but, for all prime $p \equiv 1 (\mathrm{mod}\ 4)$, we show that $\theta(Q_p) = \sqrt{p}$. Then we consider three special graphs: the Peterson graph, the icosahedron graph, and the dodecahedron graph. Also we consider an especially interesting regular graph on 7 vertices.

As out final example in Section IV, we show that Delsarte's (Ref. 3) linear programming bound for cliques in association schemes follows as a special case of our results.

II. The Lovasz Upper Bounds

In this section and the next, G will denote a fixed graph. We will continually be dealing with vectors and matrices whose components are indexed by the vertex set V of G. If $\mathbf{x}$ is such a vector, and $v \in V$, the v-th component of $\mathbf{x}$ will be denoted by $\mathbf{x}(v)$; if $\mathbf{A}$ is such a matrix, its (v, v')-th component will be denoted by $\mathbf{A}(v,v')$.

We will also be working with the *quadratic forms* associated with such vectors and matrices. If $\mathbf{x}$ is a (column) vector, and $\mathbf{A}$ a symmetric matrix, the quadratic form $\mathbf{x}^T\mathbf{A}\mathbf{x}$ is defined by

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = \sum_{(v,v')\in V^2} \mathbf{x}(v)\, \mathbf{x}(v')\, \mathbf{A}(v, v') \tag{6}$$

We will always view $\mathbf{x}^T\mathbf{A}\mathbf{x}$ as a function of the components of the vector $\mathbf{x}$.

The following will no doubt appear quite trivial, and yet it is our main result. The remainder of the paper will be devoted merely to exploring its consequences.

Theorem 1:

Let $\mathbf{A}$ be a symmetric real matrix such that

$$\begin{aligned} \mathbf{A}(v, v') &= 1 \text{ if } v = v' \\ &\leqslant 0 \text{ if } \{v, v'\} \notin G \end{aligned}$$

Then if $\mathbf{u} = (1, 1, \cdots, 1)$ denotes the all-ones vector,

$$\inf \{\mathbf{x}^T\mathbf{A}\mathbf{x} : \mathbf{x} \cdot \mathbf{u} = 1\} \leqslant \alpha(G)^{-1} \tag{7}$$

Proof:

Let $Y \subseteq V$ be a maximal independent set in G, i.e., $|Y| = \alpha(G)$. Define the vector $\mathbf{y}$ by

$$\begin{aligned} \mathbf{y}(v) &= \alpha(G)^{-1} \text{ if } v \in Y \\ &= 0 \qquad \text{if not} \end{aligned}$$

Then clearly $\mathbf{y} \cdot \mathbf{u} = 1$, and by Eq. (6), $\mathbf{y}^T\mathbf{A}\mathbf{y} \leqslant \alpha(G)^{-1}$. This proves Eq. (7).

Let us denote by $\lambda(\mathbf{A})$ the value of the left side of Eq. (7)

$$\lambda(\mathbf{A}) = \inf \{\mathbf{x}^T\mathbf{A}\mathbf{x} : \mathbf{x} \cdot \mathbf{u} = 1\} \tag{8}$$

Theorem 1 gives an upper bound on $\alpha(G)$, viz., $\alpha(G) \leqslant \lambda(\mathbf{A})^{-1}$ provided $\lambda(\mathbf{A}) > 0$. Clearly in order to apply this bound we will need to know more about the function $\lambda(\mathbf{A})$. For future reference we now list some of its more important properties. Throughout we assume $\mathbf{A}$ is real and symmetric. (Proofs of these facts may be found in Appendix B.)

First of all, unless $\mathbf{A}$ is *positive semi-definite* (hereafter abbreviated p.s.d.), i.e., unless $\mathbf{x}^T\mathbf{A}\mathbf{x} \geqslant 0$ for all $\mathbf{x}$, $\lambda(\mathbf{A})$ is negative:

$$\lambda(\mathbf{A}) \geqslant 0 \text{ iff } \mathbf{A} \text{ is p.s.d.} \tag{9}$$

Assuming that $\mathbf{A}$ is p.s.d., let $\xi_1, \cdots, \xi_N$ be a complete orthonormal set of eigenvectors for $\mathbf{A}$, i.e., $\xi_i \cdot \xi_j = \delta_{i,j}$, and $\mathbf{A}\xi_j = \lambda_j\, \xi_j$. Since $\mathbf{A}$ is p.s.d. the eigenvalues λ_j are all nonnegative. Let $\mathbf{u} = u_1\, \xi_1 + \cdots + u_N\, \xi_N$ be the expansion of $\mathbf{u}$ with respect to this basis.

Then

$$\lambda(\mathbf{A}) = 0 \text{ if } u_j \neq 0,\ \ \lambda_j = 0 \text{ for some } j \tag{10}$$

$$\lambda(\mathbf{A}) = (\Sigma\, u_j^2/\lambda_j)^{-1} \text{ otherwise} \tag{11}$$

where the summation in Eq. (11) is extended only over subscripts j with $\lambda_j > 0$. (Alternatively Eq. (11) includes Eq. (10) as a special case if we extend the summation over all j and make the conventions that $u^2/0 = 0$ if $u = 0$, $= \infty$ if $u \neq 0$, and $\infty^{-1} = 0$.)

If $\mathbf{u}$ is itself an eigenvector of $\mathbf{A}$ with eigenvalue σ, the computation of $\lambda(\mathbf{A})$ is much simpler:

$$\lambda(\mathbf{A}) = \sigma/N \text{ if } \mathbf{Au} = \sigma\mathbf{u}, \text{ and } \mathbf{A} \text{ is p.s.d.} \tag{12}$$

There is a useful dual formulation of the definition (8) for p.s.d. matrices:

$$\lambda(\mathbf{A}) = \max \{\lambda : \mathbf{A} - \lambda\, \mathbf{J} \text{ is p.s.d.}\} \tag{13}$$

where in Eq. (13) $\mathbf{J}$ denotes the matrix of all ones.

Our last auxiliary result about $\lambda(\mathbf{A})$ is that it is *multiplicative* for p.s.d. matrices:

$$\lambda(\mathbf{A} \times \mathbf{B}) = \lambda(\mathbf{A})\, \lambda(\mathbf{B}) \text{ if } \mathbf{A} \text{ and } \mathbf{B} \text{ are p.s.d.} \tag{14}$$

where $\mathbf{A} \times \mathbf{B}$ denotes the direct product of $\mathbf{A}$ and $\mathbf{B}$.

According to Eq. (9) Theorem 1 will only give nontrivial information about $\alpha(G)$ if $\mathbf{A}$ is p.s.d. This leads us to define the following two sets of matrices.

Definition 1:

The set $\Omega(G)$ is defined as the set of all matrices $\mathbf{A} = (\mathrm{A}(v,v'))$ indexed by the vertices of G, satisfying

$$\mathbf{A} \text{ is p.s.d.} \tag{15}$$

$$\mathrm{A}(v, v) = 1 \text{ for all } v \in V \tag{16}$$

$$\mathrm{A}(v, v') \leqslant 0 \text{ if } \{v, v'\} \notin G \tag{17}$$

Definition 2:

Similarly $\Omega_0(G)$ is the set of matrices satisfying Eqs. (15) and (16), with the condition (18) replaced with the stronger condition

$$\mathrm{A}(v, v') = 0 \text{ if } \{v,v'\} \notin G \tag{18}$$

The significance of the class $\Omega(G)$ is obvious, in view of Theorem 1 and Eq. (9).

Theorem 2:

$$\alpha(G) \leqslant \lambda(\mathbf{A})^{-1} \text{ for all } \mathbf{A} \in \Omega(G)$$

The significance of $\Omega_0(G)$ is given in the next theorem, which is essentially equivalent to the Lovasz bound Eq. (5). (For a proof of this equivalence, see Appendix A.)

Theorem 3:

$$\theta(G) \leqslant \lambda(\mathbf{A})^{-1} \quad \text{for all } \mathbf{A} \in \Omega_0(G)$$

Proof:

The key to the proof is the fact that if $\mathbf{A} \in \Omega_0(G)$, then the the n-th direct power $\mathbf{A}^{[n]} = \mathbf{A} \times \mathbf{A} \times \cdots \times \mathbf{A}$ (n factors) will belong to $\Omega_0(G^n)$.

To see this, observe first that $\mathbf{A}^{[n]}$ is p.s.d., being a direct product of p.s.d. matrices. Next, if $\mathbf{v} = (v_1, \cdots, v_n)$ and $\mathbf{v}' = (v_1', \cdots, v_n')$ are vertices in G^n, then by definition of the direct product,

$$\mathrm{A}^{[n]}(\mathbf{v}, \mathbf{v}') = \prod_{j=1}^{n} \mathrm{A}(v_j, v_j') \tag{19}$$

It follows immediately that $\mathrm{A}^{[n]}(\mathbf{v}, \mathbf{v}) = 1$ for all $\mathbf{v} \in G^n$, since each of the factors on the right side of Eq. (19) will then be 1. If $\{\mathbf{v}, \mathbf{v}'\} \notin G^n$, there must exist at least one index j such that $\{v_j, v_j'\} \notin G$. Since $\mathbf{A} \in \Omega_0(G)$, it follows from Eq. (18) that $\mathrm{A}(v_j, v_j') = 0$ and hence that $\mathrm{A}^{[n]}(\mathbf{v}, \mathbf{v}') = 0$ as well. Hence $\mathbf{A}^{[n]}$ satisfies Eqs. (16), (17), and (18), and thus lies in $\Omega_0(G^n)$.

Since $\Omega_0(G^n) \subseteq \Omega(G^n)$, we may apply Theorem 2 to the matrix $\mathbf{A}^{[n]}$, and conclude that $\alpha(G^n) \leqslant \lambda(\mathbf{A}^{[n]})^{-1}$. But from Eq. (14), $\lambda(\mathbf{A}^{[n]}) = \lambda(\mathbf{A})^n$. Hence $\alpha(G^n) \leqslant \lambda(\mathbf{A})^{-n}$ for all n, and so from the definition Eq. (2) of $\theta(G)$, we get $\theta(G) \leqslant \lambda(\mathbf{A})^{-1}$.

Motivated by Theorems 2 and 3, we now define the *Lovasz bounds* $\alpha_L(G)$ and $\theta_L(G)$:

$$\alpha_L(G) = \min \{\lambda(\mathbf{A})^{-1} : \mathbf{A} \in \Omega(G)\} \tag{20}$$

$$\theta_L(G) = \min \{\lambda(\mathbf{A})^{-1} : \mathbf{A} \in \Omega_0(G)\} \tag{21}$$

We have shown in Theorems 2 and 3 that $\alpha(G) \leqslant \alpha_L(G)$, $\theta(G) \leqslant \theta_L(G)$. Unfortunately, we know of no efficient algorithm for computing $\alpha_L(G)$ and $\theta_L(G)$ for an arbitrary graph. However, we will now show that one can use the symmetries of G to simplify the calculations somewhat. In Section III we will extend these ideas and show that if G is highly symmetric, the bounds $\alpha_L(G)$ and $\theta_L(G)$ can be computed via linear programming.

A *symmetry* of the graph G is a permutation of the vertex set V that leaves the edge set E invariant. Thus if π is a permutation of V, it is a symmetry of G if and only if $\{\pi(v), \pi(v')\} \in E$ whenever $\{v, v'\} \in E$. Notice that a symmetry of G is also a symmetry of the *complementary graph* G', which has vertex set V and edge set E', the set of pairs not in E.

Let P be the group of symmetries of G, and let $E_1, \cdots, E_s$ be the orbits of E under the action of P. Similarly let $E'_1, \cdots, E'_t$ be the orbits of E'. We shall call two edges lying in the same orbit *equivalent edges*.

Now suppose that $\mathbf{A} \in \Omega(G)$ or $\Omega_0(G)$. Then it is easy to see that the matrix $\bar{\mathbf{A}}$ defined by

$$\bar{\mathbf{A}}(v, v') = \frac{1}{|P|} \sum_{\pi \in P} \mathbf{A}(\pi(v), \pi(v')) \tag{22}$$

also lies in the same set. Moreover, the matrix $\bar{\mathbf{A}}$ has the property that if $\{v_1, v'_1\}$ and $\{v_2, v'_2\}$ are equivalent edges, then $\bar{\mathbf{A}}(v_1, v'_1) = \bar{\mathbf{A}}(v_2, v'_2)$. Additionally, we can show that $\lambda(\bar{\mathbf{A}}) \geqslant \lambda(\mathbf{A})$. For if we denote by $\pi(\mathbf{A})$ the matrix with entries $\mathbf{A}(\pi(v), \pi(v'))$, then for any value of λ,

$$\bar{\mathbf{A}} - \lambda \mathbf{J} = \frac{1}{|P|} \sum_{\pi \in P} (\pi(\mathbf{A}) - \lambda \mathbf{J})$$

$$= \frac{1}{|P|} \sum_{\pi \in P} \pi(\mathbf{A} - \lambda \mathbf{J})$$

If we let $\lambda = \lambda(\mathbf{A})$ then by Eq. (13) $\mathbf{A} - \lambda\mathbf{J}$ is p.s.d. and hence so is each $\pi(\mathbf{A} - \lambda\mathbf{J})$. Thus $\bar{\mathbf{A}} - \lambda\mathbf{J}$ is p.s.d. and so by Eq. (13) $\lambda(\bar{\mathbf{A}}) \geqslant \lambda(\mathbf{A})$.

Let us denote by $\mathbf{B}_j$, $j = 1, 2, \cdots, s$ the edge incidence matrices for the edge orbits E_j:

$$\mathbf{B}_j(v, v') = 1 \text{ if } \{v, v'\} \in E_j$$

$$= 0 \text{ if not}$$

Similarly we define the matrices $\mathbf{B}'_k$, $k = 1, 2, \cdots, t$ as the edge incidence matrices for the edge orbits E'_k. Then, according to the preceding discussion, the matrix $\bar{\mathbf{A}}$ can be expressed as a linear combination of these matrices, together with the $N \times N$ identity matrix $\mathbf{I}$:

$$\bar{\mathbf{A}} = \mathbf{I} + \sum_{j=1}^{s} \mu_j \mathbf{B}_j + \sum_{k=1}^{t} \mu'_k \mathbf{B}'_k \tag{23}$$

We have thus shown that starting with any matrix $\mathbf{A}$ in $\Omega(G)$ (resp. $\Omega_0(G)$), we can construct a matrix $\bar{\mathbf{A}}$ of the form Eq. (23) lying in the same class such that $\lambda(\bar{\mathbf{A}}) \geqslant \lambda(\mathbf{A})$. What this means is that in the computation of the bounds $\alpha_L(G)$ and $\theta_L(G)$, we can safely restrict ourselves to matrices of the form of Eq. (23). More formally, we define

$$\overline{\Omega}(G) = \text{p.s.d. matrices of the form of Eq. (23)}$$

$$\text{with } \mu'_k \leqslant 0 \text{ for } k = 1, 2, \cdots, t$$

$$\overline{\Omega}_0(G) = \text{p.s.d. matrices of the form of Eq. (23)}$$

$$\text{with } \mu'_k = 0 \text{ for } k = 1, 2, \cdots, t$$

We have then the following computationally simpler definition of the Lovasz bounds:

$$\alpha_L(G) = \min \{\lambda(\mathbf{A})^{-1} : \mathbf{A} \in \overline{\Omega}(G)\}$$

$$\theta_L(G) = \min \{\lambda(\mathbf{A})^{-1} : \mathbf{A} \in \overline{\Omega}_0(G)\}$$

III. A Linear Programming Bound for α (G) and θ (G)

In this section we will show that if the graph G is sufficiently symmetric, the computation of the bounds $\alpha_L(G)$ and $\theta_L(G)$ can be greatly simplified.

The degree of symmetry we require is that the *incidence matrices* $\{\mathbf{B}_j\}$, $\{\mathbf{B}'_k\}$ in Eq. (23) *commute with each other*. That this is in fact a statement about the symmetry group of G can be seen as follows.

Suppose P is the symmetry group of G. With each $\pi \in P$ we associate the corresponding permutation matrix $\boldsymbol{\pi}^*$:

$$\boldsymbol{\pi}^*(v, v') = 1 \text{ if } \pi(v) = v'$$

$$= 0 \text{ if not}$$

Naturally the edge orbits $\{E_j\}$, $\{E'_k\}$ are left invariant by the symmetries $\pi \in P$; in terms of the corresponding incidence matrices, this can be expressed as

$$\boldsymbol{\pi}^*\mathbf{B}_j = \mathbf{B}_j\boldsymbol{\pi}^*, \quad \text{all } j = 1, 2, \cdots, s, \pi \in P \tag{24}$$

$$\boldsymbol{\pi}^*\mathbf{B}'_k = \mathbf{B}'_k\boldsymbol{\pi}^*, \text{ all } k = 1, 2, \cdots, t, \pi \in P$$

Now let P^* denote the group of all permutation matrices corresponding to the permutations in P, and let $Z(P^*)$ be the *centralizer ring* of P^*, i.e., the set of all matrices that commute with all $\boldsymbol{\pi}^* \in P^*$. According to Eq. (24), the matrices $\{\mathbf{B}_j\}$, $\{\mathbf{B}'_j\}$ all belong to $Z(P^*)$.

If the ring $Z(P^*)$ were known to be commutative, then it would follow immediately that the matrices $\mathbf{B}_j$, $\mathbf{B}'_k$ commute with each other. Fortunately, this frequently turns out to be the case. Indeed it can be shown (Ref. 10, Chapter 5) that if P is transitive, then $Z(P^*)$ is commutative if and only if the complex representation of P afforded by the matrix group P^* decomposes into a sum of inequivalent irreducible representations. In particular, if P contains a transitive abelian subgroup, or if for any pair (v, v') of distinct vertices there is an element of P that exchanges v and v', this condition will be satisfied.

Motivated by the preceding discussion, we now place our results in the following general setting.

Let V be a finite set containing N elements, and let $\{E_1, E_2, \cdots, E_n\}$ be a partition of the collection E of all two-element subsets of V. For each $j = 1, 2, \cdots, n$ let $\mathbf{A}_j$ be the incidence matrix for E_j:

$$\mathbf{A}_j(v, v') = 1 \text{ if } \{v, v'\} \in E_j$$

$$= 0 \text{ if not}$$

Let $\mathbf{A}_0$ denote the $N \times N$ identity matrix. Assume that the matrices $\{\mathbf{A}_j : j = 0, 1, \cdots, n\}$ commute with each other. In summary, the assumptions are that the $\mathbf{A}_j$'s are (0, 1) matrices satisfying

$$\mathbf{A}_0 = I, \sum_{j=0}^{n} \mathbf{A}_j = J \tag{25}$$

$$\text{Each } \mathbf{A}_j \text{ is symmetric} \tag{26}$$

$$\mathbf{A}_j\mathbf{A}_k = \mathbf{A}_k\mathbf{A}_j \quad \text{for all } j, k = 0, 1, \cdots, n \tag{27}$$

If C is a fixed subset of $\{1, 2, \cdots, n\}$, let G_C be the graph with vertex set V and edge set

$$E_C = \bigcup_{j \in C} E_j \tag{28}$$

Our goal is to give "linear programming" upper bounds on $\alpha(G_C)$ and $\theta(G_C)$ (Theorems 4 and 5, below). To state these results, however, we need some preliminary discussion.

Notice that because of Eqs. (25) and (27), each matrix $\mathbf{A}_j$ commutes with $\mathbf{J}$, the all ones matrix, and hence the $n + 2$ matrices $\mathbf{J}$, $\mathbf{A}_0, \cdots, \mathbf{A}_n$ all commute with each other. Since these matrices are moreover symmetric, and hence diagonalizable, it follows from a known theorem of linear algebra (see Ref. 4, Chapter 6, Theorem 4), that there exists a set $\{\xi_m\}_{m=1}^{N}$ of linearly independent *simultaneous eigenvectors* for these $n + 2$ matrices.

In particular the ξ_m's are eigenvectors for J:

$$\mathbf{J}\xi_m = \lambda_m \xi_m, \; m = 1, 2, \cdots, N \tag{29}$$

$\{\lambda_m\}$ being the set of eigenvalues for $\mathbf{J}$. But $\mathbf{J}$ has only the eigenvalues $\{0, N\}$, and a simple calculation shows that if $\mathbf{J}\xi = N\xi$, then ξ must be a scalar multiple of $\mathbf{u}$, the all ones vector. Thus we may assume that $\xi_1 = \mathbf{u}$.

Now for each j, m, define the eigenvalues $\lambda_{j,m}$ by

$$\mathbf{A}_j \xi_m = \lambda_{j,m} \xi_m \quad j = 0, 1, \cdots, n \tag{30}$$

$$m = 1, 2, \cdots, N$$

We come now to our "linear programming" bounds for $\alpha(G)$ and $\theta(G)$.

Theorem 4:

Let $\mu_1, \mu_2, \cdots, \mu_n$ be real numbers such that

$$\mu_j \leqslant 0 \quad \text{if } j \notin C \tag{31}$$

and

$$1 + \sum_{j=1}^{n} \mu_j \lambda_{j,m} \geqslant 0, \quad m = 1, 2, \cdots, N \tag{32}$$

Then

$$\alpha(G_C) \leqslant N \bigg/ \left(1 + \sum_{j=1}^{n} \mu_j \lambda_{j,1}\right)$$

Theorem 5:

Let $\mu_1, \cdots, \mu_n$ satisfy Eq. (32), and also

$$\mu_j = 0 \text{ if } j \notin C \tag{33}$$

Then

$$\theta(G_C) \leqslant N \bigg/ \left(1 + \sum_{j=1}^{n} \mu_j \lambda_{j,1}\right)$$

Proofs:

For the given constants $\{\mu_j\}$, define

$$\mathbf{A} = \mathbf{I} + \sum_{j=1}^{n} \mu_j \mathbf{A}_j$$

Clearly the vectors $\{\boldsymbol{\xi}_m\}$ are eigenvectors for $\mathbf{A}$, since

$$\mathbf{A}\,\boldsymbol{\xi}_m = \boldsymbol{\xi}_m + \sum_{j=1}^{n} \mu_j \mathbf{A}_j \boldsymbol{\xi}_m$$

$$= \left(1 + \sum_{j=1}^{m} \mu_j \lambda_{j,m}\right) \boldsymbol{\xi}_m$$

Furthermore, the hypothesis Eq. (32) ensures that the eigenvalues $\{1 + \Sigma\, \mu_j \lambda_{j,m}\}$ of $\mathbf{A}$ are nonegative, and hence that $\mathbf{A}$ is p.s.d. The conditions (31) and (33) now imply that the matrix $\mathbf{A}$ belongs to $\Omega\,(G_C)$ or $\Omega_0(G_C)$. Thus by Theorems 2 and 3, we get $\alpha(G_C) \leqslant \lambda(\mathbf{A})^{-1}$, $\theta(G_C) \leqslant \lambda(\mathbf{A})^{-1}$.

To compute $\lambda(\mathbf{A})$ observe that $\boldsymbol{\xi}_1 = \mathbf{u}$ is an eigenvector for $\mathbf{A}$, with corresponding eigenvalue

$$1 + \sum_{j=1}^{n} \mu_j \lambda_{j,1}$$

and so by Eq. (12), $\lambda(\mathbf{A}) = (1 + \Sigma\, \mu_j \lambda_{j,1})/N$. Theorems 4 and 5 now follow.

To get the best possible bounds of the kind given in Theorems 4 and 5, we are essentially required to maximize the linear function

$$\sum_{1}^{n} \mu_j \lambda_{j,1}$$

subject to the linear constraints (31) or (33), and (32). This is a linear programming problem (once the eigenvectors and eigenvalues of the matrices $\mathbf{A}_j$ are known); and hence we have succeeded in showing that *the Lovasz bounds $\alpha_L(G)$ and $\theta_L(G)$ can be computed via linear programming, provided the incidence matrices $\mathbf{B}_j$, $\mathbf{B}'_k$ of the edge orbits E_j, E'_k commute.*

IV. Some Applications

In this section we will describe a few of the many possible applications of the preceding results. In particular we will obtain Lovasz' original result on the capacity of C_5, and Delsarte's linear programming bound for cliques in association schemes.

A. Example 1. Regular Graphs

A graph G is said to be *regular* if the number of edges containing a given vertex v is a constant r, independent of v, called the *valence* of G.

A consequence of the regularity of G is that the incidence matrix $\mathbf{B}$ corresponding to the edge set E commutes with J: $\mathbf{JB} = \mathbf{BJ}$. Obviously $\mathbf{B}$ also commutes with the identity matrix $\mathbf{I}$, and hence also with $\mathbf{B}' = \mathbf{J} - \mathbf{I} - \mathbf{B}$, which is the edge incidence matrix of the complementary graph. We are thus in a position to apply Theorem 5. Omitting the straightforward details, the result is

$$\theta(G) \leqslant \frac{N}{1 + r/|\lambda_{\min}|} \tag{34}$$

where $\lambda_{\min}$ is the smallest eigenvalue of $\mathbf{B}$ (which is necessarily negative, unless E is empty.)

Furthermore, if in addition the group of symmetries of G permutes the edges transitively, it follows from our results that $\theta_L(G)$ is equal to the right side of Eq. (34), i.e., Eq. (34) is the best possible bound of this type. (The bound Eq. (34) is equivalent to Lovasz' Theorem 9.)

B. Example 2. The Graphs C_N

Denote by C_N the cyclic graph on N vertices, i.e, $V = \{0, 1, \cdots, N-1\}$, $E = \{\{i, i+1\} : i = 0, 1, \cdots, N-1\}$, with indices taken mod. N. These graphs are all regular, and indeed the cyclic group of order N permutes the edges transitively, so we may compute $\theta_L(C_N)$ by the formula (34).

To find the eigenvalues of the incidence matrix $\mathbf{B}$ in this case, observe that the vectors $\mathbf{x}(\zeta) = (1, \zeta, \cdots, \zeta^{N-1})$, where ζ is any complex N-th root of unity, form an independent set of eigenvectors for $\mathbf{B}$, and indeed

$$\mathbf{B}\mathbf{x}(\zeta) = (\zeta + \zeta^{-1})\mathbf{x}(\zeta)$$

Hence the eigenvalues of $\mathbf{B}$ are $\{(\zeta + \zeta^{-1})\} = \{2\cos(2\pi k/N) : k = 0, 1, \cdots, \lfloor N/2 \rfloor\}$. The least member of this set is clearly -2 if N is even, and $2\cos(\pi - \pi/N) = -2\cos(\pi/N)$ if N is odd and $\geqslant 3$. Thus by Eq. (34).

$$\begin{aligned} \theta(C_N) &\leqslant N/2, N \text{ even} \\ &\leqslant N/(1 + (\cos \pi/N)^{-1}), \quad N \text{ odd}, \geqslant 3 \end{aligned} \tag{35}$$

For N even, or $N = 3$, this bound is sharp; but these results are quite elementary and were already known to Shannon.

For odd $N \geqslant 5$, however, the bounds are nontrivial. (They appear as Corollary 5 in Lovasz.) With $N = 5$, for example, we have $\theta(C_5) \leqslant \sqrt{5}$.

For odd $N \geqslant 7$, the upper and lower bounds on $\theta(C_N)$ do not agree. Here is a table of the upper bound (35) vs the best known lower bounds (Ref. 1) for odd $7 \leqslant N \leqslant 19$:

$$7^{3/5} = 3.21410 \leqslant \theta(C_7) \leqslant 3.31767$$

$$81^{1/3} = 4.32675 \leqslant \theta(C_9) \leqslant 4.36009$$

$$148^{1/3} = 5.28957 \leqslant \theta(C_{11}) \leqslant 5.38630$$

$$247^{1/3} = 6.27431 \leqslant \theta(C_{13}) \leqslant 6.40417$$

$$380^{1/3} = 7.24316 \leqslant \theta(C_{15}) \leqslant 7.41715$$

$$4913^{1/4} = 8.37214 \leqslant \theta(C_{17}) \leqslant 8.42701$$

$$7666^{1/4} = 9.35712 \leqslant \theta(C_{19}) \leqslant 9.43477$$

C. Example 3. The Quadratic Residue Graphs

Let $p \equiv 1 \pmod 4$ be a prime. The graph Q_p has vertex set $V = \{0, 1, \cdots, p-1\}$, and edge set $E = \{\{v, v'\}: v - v' \text{ is a quadratic residue (q.r.) mod } p\}$. (Note that Q_5 is isomorphic to the pentagonal graph C_5.) Q_p is regular with valence $(p-1)/2$. The edge-incidence matrix is given by

$$\begin{aligned} \mathbf{B}_p(v - v') &= 1 \text{ if } v - v' \text{ is a q.r. } (\text{mod } p) \\ &\quad 0 \text{ if not} \end{aligned}$$

One easily verifies that the p vectors $\mathbf{x}(\xi) = (1, \xi, \cdots, \xi^{p-1})$, where ξ is any complex p-th root of unity, are eigenvectors for $\mathbf{B}_p$, and that the eigenvalue associated with $\mathbf{x}(\xi)$ is $\Sigma\ \{\xi^a : a \text{ is a q.r.}\}$. It is well known (see Ref. 9, Section 11, for example) that these sums assume only the three distinct values $(p-1)/2$, $(-1 \pm \sqrt{p})/2$. Hence the least eigenvalue of $\mathbf{B}_p$ is $(-1 - \sqrt{p})/2$, and Eq. (34) yields $\theta(Q_p) \leqslant \sqrt{p}$. On the other hand, if b is a fixed quadratic nonresidue (mod p), the p ordered pairs (v, bv), $v \in V$ form an independent set in Q_p^2, and hence $\alpha(Q_p^2) \geqslant p$. These two inequalities establish the fact that $\theta(Q_p) = \sqrt{p}$, for all $p \equiv 1 \pmod 4$. Because of this result, it is clear that the graphs Q_p form a more satisfactory generalization of the pentagon of Fig. 1 than the graphs C_N. (These graphs do not appear in Lovasz' paper. But he does show in Theorem 12 that if G is self-complementary, and if the symmetry group of G is transitive on the vertics, then $\theta(G) = \sqrt{|V|}$. This example is thus an explicit case of Lovasz' Theorem 12.)

D. Example 4. Some Miscellaneous Edge-Transitive Graphs

Here we will apply the bound (34) to three particularly interesting graphs. In each case there is only one equivalence class of edges, so that the bounds obtained are all equal to $\theta_L(G)$. In each case $\theta_L(G)$ is strictly less than any bound that could be obtained by Shannon's techniques.

The Peterson Graph (Fig. 2) is a regular graph with $N = 10$, $r = 3$. The minimum eigenvalue here turns out to be -2, so (34) yields $\theta(G) \leqslant 4$. On the other hand $\alpha(G) = 4$ (note the four circled vertices in Fig. 2), and so $\theta(G) = 4$. (This result is a special case of Lovasz' Theorem 13.)

The Icosahedron Graph (Fig. 3) has $N = 12$, $r = 5$; its vertices and edges are formed from those of the regular icosahedron. Here the minimum eigenvalue is $-\sqrt{5}$, and so from (34), $\theta(G) \leqslant 3(\sqrt{5} - 1) = 3.7082$. On the other hand $\alpha(G) = 3$, so we have $3 \leqslant \theta(G) \leqslant 3.7082$.

The Dodecahedron Graph (Fig. 4) is the graph of the regular dodecahedron, with $N = 20$, $r = 3$. Here $\lambda_{\min} = -\sqrt{5}$ also; hence (34) gives $\theta(G) \leqslant 15\sqrt{5} - 25 = 8.5410$. On the other hand, $\alpha(G) = 8$, as shown. Thus $8 \leqslant \theta(G) \leqslant 8.5410$.

E. Example 5. A Special Graph on 7 Vertices

Consider the graph depicted in Fig. 5. This graph is regular with $N = 7$, $r = 4$, and minimum eigenvalue $= 4 \cos 6\pi/7 \cos 2\pi/7 = -2.2470$, and hence by (34), $\theta(G) \leqslant 2.5178$. However, under the action of the symmetries of G, there are two equivalence classes of edges: those of type $\{i, i+1\}$, and those of type $\{i, i+2\}$, modulo 7. In this case the bound of (34) is strictly larger than the Lovasz bound $\theta_L(G)$; in order to compute $\theta_L(G)$, we must apply Theorem 5 directly.

Thus let $\mathbf{A}_0$ denote the 7×7 identity matrix: $\mathbf{A}_1$, the incidence matrix for edges of type $\{i, i+1\}$; $\mathbf{A}_2$, for edges of type $\{i, i+2\}$: and $\mathbf{A}_3$, for edges of type $\{i, i+3\}$ (which are not edges of G.)

One easily verifies that the matrices $\{\mathbf{A}_0, \mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3\}$ satisfy conditions (25) through (27). If $C = \{1,2\}$ the graph G_C is the graph of Fig. 5. Also, the 7 vectors of the form $x(\zeta) = (1, \zeta, \cdots, \zeta^6)$, where ζ is a complex 7-th root of unity, form a set of common eigenvectors for the A's:

$$\mathbf{A}_0\, \mathbf{x}(\zeta) = \mathbf{x}(\zeta)$$

$$\mathbf{A}_1\, \mathbf{x}(\zeta) = (\zeta + \zeta^{-1})\, \mathbf{x}(\zeta)$$

$$\mathbf{A}_2\, \mathbf{x}(\zeta) = (\zeta^2 + \zeta^{-2})\, \mathbf{x}(\zeta)$$

$$\mathbf{A}_3\, \mathbf{x}(\zeta) = (\zeta^3 + \zeta^{-3})\, \mathbf{x}(\zeta)$$

Thus according to Theorem 5, if μ_1 and μ_2 satisfy $1 + \mu_1(\zeta + \zeta^{-1}) + \mu_2(\zeta^2 + \zeta^{-2}) \geqslant 0$ for all 7-th roots of unity ζ, then $\theta(G) \leqslant 7/(1 + 2\mu_1 + 2\mu_2)$. To get the best possible such bound we must maximize the function $\mu_1 + \mu_2$ subject to the above set of inequalities. This is easily done by hand[4], and we get the largest possible value with $\mu_1 = 0.8020$, $\mu_2 = 0.3569$. The resulting bound is $\theta(G) \leqslant 2.1098$. (This graph is the complement of the cyclic graph C_7; we denote this by writing $G = C_7'$. Lovasz' Theorem 8 asserts that if G is any graph with a vertex-transitive symmetry group, then $\theta_L(G)\, \theta_L(G') = N$. It thus follows from Example 2 that $\theta_L(C_N') = 1 + (\cos \pi/N)^{-1}$ for odd N. We have included this alternate derivation only to illustrate a nontrivial example of our linear programming approach.)

F. Example 6. Delsarte's Linear Programming Bounds

Delsarte (Ref. 3) obtained linear programming bounds for cliques in association schemes. Here we sketch a demonstration that these bounds are subsumed under our Theorem 4. (Recently Schrijver, (Ref. 7), has obtained similar results.)

Let V be a finite set, and let $R = \{R_0, R_1, \cdots, R_n\}$ be a family of $n + 1$ subsets of the Cartesian square V^2. The R_j are *relations* on V and can be described by their incidence matrices

$$\mathbf{A}_j(\nu, \nu') = 1 \text{ if } (\nu, \nu') \in R_j$$

$$= 0 \text{ if not}$$

[4] There are really only three inequalities to consider, viz. those with $\zeta = \exp(2\pi i k/7)$, $k = 1, 2, 3$.

The pair (V, R) is called a (symmetric) *association scheme* if the following conditions are satisfied:

(1) R is a partition of V^2, and R_0 is the diagonal, i.e., $R_0 = \{(v, v) : v \in V\}$.

(2) The relations $\{R_j\}$ are symmetric, i.e., $(v, v') \in R_j$ implies $(v', v) \in R_j$.

(3) There exist numbers $p_{i,j}^{(k)} = p_{j,i}^{(k)}$ such that for all $i, j = 0, 1, \cdots n$,

$$\mathbf{A}_i \mathbf{A}_j = \sum_{k=0}^{n} p_{i,j}^{(k)} \mathbf{A}_k$$

If M is a subset of $\{0, 1, \cdots, n\}$ with $0 \in M$, a nonempty subset $Y \subseteq V$ is called an *M-clique with respect to R* if it satisfies

$$R_j \cap Y^2 = \phi \text{ for all } j \notin M$$

Delsarte (Ref. 3) gives an upper bound on the number of points in an M-clique, which is the value of a certain linear program.

But we can equally apply our Theorem 4 to the same problem, for the matrices $\{\mathbf{A}_j\}$ of the association scheme certainly satisfy conditions (25) through (27) (note that since $p_{i,j}^{(k)} = p_{j,i}^{(k)}$, condition (3) is considerably stronger than (27).) If we let $C = \{j : j \notin M\}$, then an M-clique as defined above is an independent set in G_C, and so the upper bound of Theorem 4 is an upper bound on the cardinality of any M-clique in G_C. One can in fact show that this is the same bound as Delsarte's. Hence Theorem 4 is more general than Delsarte's bound, since it applies to many cases that are not association schemes.

Appendix A

Equivalence of Theorem 3 with Lovasz' Bound

Given an orthonormal representation of the graph G, define the matrix $\mathbf{A}$ by

$$\mathbf{A}(\mathbf{v}, \mathbf{v}') = \mathbf{v} \cdot \mathbf{v}'$$

Clearly $\mathbf{A} \in \Omega_0(G)$. If $\mathbf{b}$ is a unit vector and if $\lambda = \min \{(\mathbf{v} \cdot \mathbf{b})^2 : \mathbf{v} \in \mathrm{V}\}$, then $\mathbf{A} - \lambda\mathbf{J}$ is p.s.d. This is because we can write $\mathbf{A} - \lambda\mathbf{J} = \mathbf{B} + \mathbf{C}$, where the matrices $\mathbf{B}$ and $\mathbf{C}$ are defined by

$$\mathbf{B}(\mathbf{v}, \mathbf{v}') = (\mathbf{v} - (\mathbf{v} \cdot \mathbf{b})\mathbf{b}) \cdot (\mathbf{v}' - (\mathbf{v}' \cdot \mathbf{b})\mathbf{b})$$

$$\mathbf{C}(\mathbf{v}, \mathbf{v}') = (\mathbf{v} \cdot \mathbf{b})(\mathbf{v}' \cdot \mathbf{b}) - \lambda$$

$\mathbf{B}$ is p.s.d., since it is the matrix of inner products of a set of vectors. $\mathbf{C}$ is also p.s.d., since if $\{\mathbf{x}(\mathbf{v}) : \mathbf{v} \in \mathrm{V}\}$ is any set of real numbers,

$$\mathbf{x}^T\mathbf{C}\mathbf{x} = (\sum_{\mathbf{v}} \mathbf{x}(\mathbf{v})(\mathbf{v} \cdot \mathbf{b}))^2 - \lambda(\Sigma\, \mathbf{x}(\mathbf{v}))^2$$

$$\geqslant 0$$

since $(\mathbf{v} \cdot \mathbf{b}) \geqslant \sqrt{\lambda}$ for all $\mathbf{v}$. Thus $\mathbf{A} - \lambda\mathbf{J}$, being the sum of two p.s.d. matrices, is also p.s.d.

Since $\mathbf{A} \in \Omega_0(G)$ and $\mathbf{A} - \lambda\mathbf{J}$ is p.s.d., it now follows from (13) that $\lambda(\mathbf{A}) \geqslant \lambda$ and hence from Theorem 3 that $\theta(G) \leqslant \lambda^{-1} = (\min \{(\mathbf{v} \cdot \mathbf{b})^2 : \mathbf{v} \in \mathrm{V}\})^{-1}$. Thus Theorem 3 implies the Lovasz bound (1.5).

Conversely, if $\mathbf{A} \in \Omega_0(G)$, let $\lambda = \lambda(\mathbf{A})$. Then $\mathbf{A} - \lambda\mathbf{J}$ is p.s.d., and from a known theorem (Ref. 2), Chapter 9), there exists a matrix $\mathbf{B}$ such that $\mathbf{B}^T\mathbf{B} = \mathbf{A} - \lambda\mathbf{J}$. Letting $\{\mathbf{w}(\nu) : \nu \in V\}$ denote the column vectors of $\mathbf{B}$, we have

$$\mathbf{w}(\nu) \cdot \mathbf{w}(\nu') = \mathbf{A}(\nu, \nu') - \lambda$$

$$= 1 - \lambda \text{ if } \nu = \nu'$$

$$= -\lambda \text{ if } \{\nu, \nu'\} \notin G.$$

Now, let $\mathbf{t}$ be a vector orthogonal to all the $\mathbf{w}(\nu)$'s with $|\mathbf{t}|^2 = \lambda$ (increase the dimension of the underlying space, if necessary), and define

$$\mathbf{x}(\nu) = \mathbf{w}(\nu) + t$$

The $\mathbf{x}(\nu)$'s are unit vectors, since

$$\mathbf{x}(\nu) \cdot \mathbf{x}(\nu') = \mathbf{A}(\nu, \nu')$$

The orthogonality graph defined by the $\mathbf{x}$'s is thus a subgraph (same vertex set, a subset of edges) G' of G.

Furthermore, if we define the unit vector $\mathbf{b}$:

$$\mathbf{b} = \frac{\mathbf{t}}{|\mathbf{t}|}$$

we have $\mathbf{x}(\nu) \cdot \mathbf{b} = |\mathbf{t}| = \sqrt{\lambda}$ for all $\nu \in V$. Hence by Lovasz' bond (5) $\theta(G') \leqslant \lambda^{-1} = \lambda(\mathbf{A})^{-1}$. But clearly $\theta(G) \leqslant \theta(G')$, and so Lovasz' result implies our Theorem 3.

Appendix B

Proof of Assertions (9) through (14)

Recall that $\mathbf{A}$ is a real symmetric $N \times N$ matrix, and that

$$\lambda(\mathbf{A}) = \inf \{\mathbf{x}^T\mathbf{A}\mathbf{x} : \mathbf{x} \cdot \mathbf{u} = 1\} \qquad \text{(B-1)}$$

where $\mathbf{u} = (1, 1, \cdots, 1)$ is the all ones vector. According to the Principal Axis Theorem, (Ref. 2), (Chapters 9 and 10), there exists a set $\{\xi_1, \cdots, \xi_N\}$ of N orthogonal eigenvectors of $\mathbf{A}$:

$$\xi_i \cdot \xi_j = 1 \text{ if } i = j \qquad \text{(B-2)}$$

$$= 0 \text{ if } i \neq j$$

$$\mathbf{A}\xi_j = \lambda_j \xi_j, \ = 1, 2, \cdots, N \qquad \text{(B-3)}$$

Thus, if $\mathbf{x} = x_1 \xi_1 + \cdots + x_N \xi_N$ and $\mathbf{y} = y_1 \xi_1 + \cdots + y_N \xi_N$, we have:

$$\mathbf{x} \cdot \mathbf{y} = \sum_{j=1}^{N} x_j y_j \qquad \text{(B-4)}$$

$$\mathbf{x}^T\mathbf{A}\mathbf{y} = \sum_{j=1}^{N} \lambda_j x_j y_j \qquad \text{(B-5)}$$

If one of the eigenvalues λ_j is negative, we can construct a vector $\mathbf{x}$ with $\mathbf{x} \cdot \mathbf{u} = 1$ and $\mathbf{x}^T\mathbf{A}\mathbf{x} < 0$, as follows. Let $\mathbf{v} = v_1 \xi_1 + \cdots + v_N \xi_N$ be a fixed vector with $\mathbf{v} \cdot \mathbf{u} = 1$, and define for any real β

$$\mathbf{x} = \frac{\beta \xi_j + \mathbf{v}}{\beta u_j + 1} \qquad \text{(B-6)}$$

where $\mathbf{u} = u_1 \xi_1 + \cdots + u_N \xi_N$ is the expansion of $\mathbf{u}$. Clearly $\mathbf{x} \cdot \mathbf{u} = 1$, and from (B-5) we compute

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = \frac{1}{(\beta u_j + 1)^2} \left\{ \lambda_j \beta^2 + 2\beta \lambda_j v_j + \sum_{i=1}^{N} \lambda_j v_j^2 \right\}$$

Clearly this will be negative if β is large enough, since the expression in brackets is then dominated by the term $-|\lambda_j| \beta^2$. This proves (9).

Thus we assume $\mathbf{A}$ is p.s.d., i.e., that the eigenvalues $\{\lambda_j\}$ are all nonnegative. If for some index j we have $\lambda_j = 0$ but $u_j \neq 0$, and if we set $\mathbf{x} = u_j^{-1} \xi_j$, then $\mathbf{x} \cdot \mathbf{u} = 1$, $\mathbf{x}^T\mathbf{A}\mathbf{x} = 0$. This proves (10).

On the other hand if $u_j = 0$ whenever $\lambda_j = 0$, we get by Schwarz' inequality

$$(\Sigma x_j u_j)^2 \leqslant (\Sigma \lambda_j x_j^2)(\Sigma \lambda_j^{-1} u_j^2) \qquad \text{(B-7)}$$

where the summation is extended only over indices for which $\lambda_j \geqslant 0$. Since by (B-4) and (B-5) $\Sigma x_j u_j = x \cdot u$, and $\Sigma \lambda_j x_j^2 = \mathbf{x}^T\mathbf{A}\mathbf{x}$, (B-7) immediately implies that if $\mathbf{x} \cdot \mathbf{u} = 1$, then $\mathbf{x}^T\mathbf{A}\mathbf{x} \geqslant (\Sigma\lambda_j^{-1} u_j^2)^{-1} = \lambda$. On the other hand, by choosing $x_i = u_i \lambda_i^{-1} \lambda$ for all i, we get $\mathbf{x} \cdot \mathbf{u} = 1$ and $\mathbf{x}^T\mathbf{A}\mathbf{x} = \lambda$. This proves (11).

If $\mathbf{u}$ is itself an eigenvector for $\mathbf{A}$, with eigenvalue σ, then in the expansion $\mathbf{u} = u_1 \xi_1 + \cdots + u_N \xi_N$, u_j must be zero unless $\lambda_j = \sigma$. Hence $\lambda(\mathbf{A}) = (\sigma^{-1} \Sigma u_j^2)^{-1} = \sigma/N$, since $\Sigma u_j^2 = \mathbf{u} \cdot \mathbf{u} = N$. This proves (12).

To prove (13), observe that

$$\mathbf{x}^T(\mathbf{A} - \lambda\mathbf{J})\mathbf{x} = \mathbf{x}^T\mathbf{A}\mathbf{x} - \lambda(\mathbf{x} \cdot \mathbf{u})^2$$

$$= \Sigma \lambda_j x_j^2 - \lambda(\Sigma x_j u_j)^2 \qquad \text{(B-8)}$$

Comparing this to (B-7), we see that this expression will be nonnegative for all x if and only if $\lambda \leqslant (\Sigma\lambda_j^{-1} u_j^2)^{-1}$, i.e. $\lambda \leqslant \lambda(\mathbf{A})$. This proves (13).

Finally we turn to (14). Suppose then that $\xi_1, \cdots, \xi_N$ are principal axes for the matrix $\mathbf{A}$, with corresponding eigenvalues $\{\lambda_j\}$, and that $\eta_1, \cdots, \eta_M$ are principal axes for $\mathbf{B}$, with eigenvalues $\{\mu_k\}$. Suppose further that $\mathbf{u}^{(A)} = u_1^{(A)} \xi_1 + \ldots + u_N^{(A)} \xi_N$, $u^{(B)} = u_1^{(B)} \eta_1 + \cdots + u_M^{(B)} \eta_M$ are the expansions of all ones vectors, with respect to these two bases.

It follows from known results (see Ref. 6, Section VII) that the MN vectors $\xi_j \times \eta_k$ are principal axes for the matrix $A \times B$, with associated eigenvalues $\lambda_j \mu_k$. Furthermore, the expansion of the MN – dimensional all-ones vector with respect to the basis $\{\xi_j \times \eta_k\}$ is clearly

$$\mathbf{u} = \sum_{j,k} u_j^{(A)} u_k^{(B)} (\xi_j \times \eta_k)$$

Thus, according to (11),

$$\lambda(\mathbf{A} \times \mathbf{B}) = \left(\sum_{j,k} (\lambda_j \mu_k)^{-1} \left[u_j^{(A)} u_k^{(B)} \right]^2 \right)^{-1}$$

$$= \lambda(\mathbf{A})\, \lambda(\mathbf{B}),$$

establishing (14).

References

1. Baumert, L. D., et. al., "A Combinatorial Packing Problem," pp. 97-108 in *Computers in Algebra and Number Theory*, SIAM-AMS Proceedings Vol. IV. Providence: American Mathematical Society, 1971.
2. Birkhoff, G., and MacLane, S., *A Survey of Modern Algebra*. New York: MacMillan, 1960.
3. Delsarte, P., "An Algebraic Approach to the Association Schemes of Coding Theory," Philips Research Reports Supplement 1973, No. 10.
4. Hoffman, K., and Kunze, R., *Linear Algebra*. Englewood Cliffs: Prentice – Hall, Inc., 1961.
5. Lovasz, L., "On the Shannon Capacity of a Graph," *IEEE Trans. Inform. Theory*, to appear.
6. MacDuffee, C., *The Theory of Matrices*. New York: Chelsea, 1956.
7. Schrijver, A., "A Comparison of the Bounds of Delsarte and Lovasz," *IEEE Trans. Inform. Theory*, to appear.
8. Shannon, C., "The Zero Error Capacity of a Noisy Channel," *IRE Trans. Inform. Theory*, Vol. 1T-2, pp. 8-19, Sept. 1956.
9. Weyl, H., *Algebraic Theory of Numbers* (Annals of Math. Studies No. 1). Princeton: Princeton University Press, 1940.
10. Wielandt, M. *Finite Permutation Groups*. New York: Academic Press, 1964.

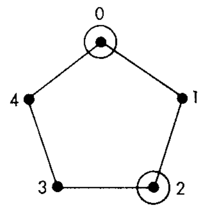

Fig. 1. The graph C_5; α (C_5) = 2 (this is also the graph Q_5 of Example 3 in Section IV)

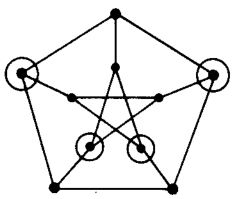

Fig. 2. The Peterson Graph

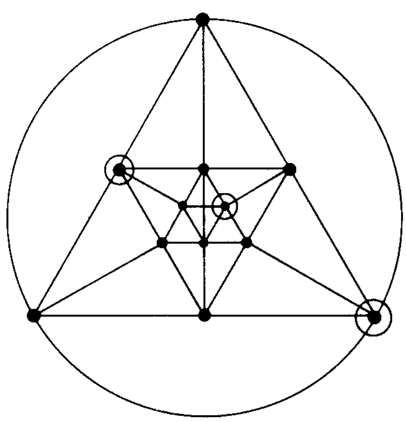

Fig. 3. The Icosahedron Graph

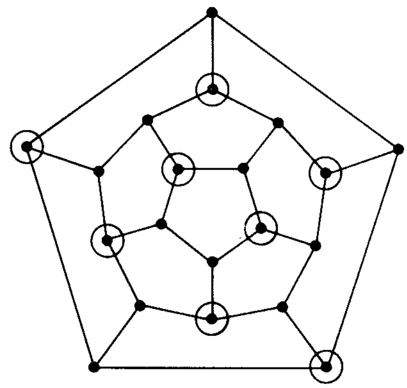

Fig. 4. The Dodecahedron Graph

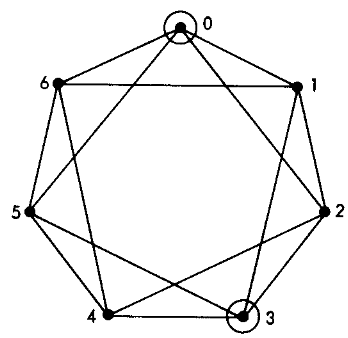

Fig. 5. A regular graph on seven vertices